

CLAIMS:

We claim:

1. An autonomic request routing policy selection system comprising:
a plurality of pre-configured request routing policies;
a data store of cache metrics for said pre-configured request routing policies;
and,
a routing policy selector configured for communicative linkage to a server cluster comprising a plurality of servers and programmed to select a particular one of said request routing policies for use in routing content requests in said server cluster based upon said cache metrics, said routing policy selector further comprising a coupling to said routing policies and said data store of cache metrics.
2. The system of claim 1, wherein said pre-configured request routing policies comprise a Layer 4 request routing policy and a Layer 7 request routing policy.
3. The system of claim 2, wherein said Layer 4 request routing policy comprises a server load balancing type policy.
4. The system of claim 2, wherein said Layer 7 request routing policy comprises a content localizing type policy.
5. The system of claim 4, wherein said content localizing type policy comprises a uniform resource locator (URL) hashing policy.

6. The system of claim 1, wherein said cache metrics comprises a plurality of Zipf-like analyses based upon different selected alpha values for different workloads imposed upon said server cluster according to different ones of said request routing policies.

7. An autonomic request routing policy selection method comprising the steps of:
identifying a contemporary trace footprint experienced by a coupled server cluster;
identifying a cache allocation for said coupled server cluster;
retrieving at least two sets of hit rate metrics, each set of metrics corresponding to a particular routing policy;
comparing said hit rate metrics based upon said identified trace footprint and said identified cache allocation to determine a preferred routing policy; and,
selecting said preferred routing policy for use in routing content requests to said server cluster.

8. The method of claim 7, further comprising the steps of:
computing with said hit rate metrics, an optimal server cluster configuration for said preferred routing policy; and,
provisioning an optimal number of servers in said server cluster based upon said computed optimal server cluster configuration.

9. The method of claim 7, wherein said selecting step comprises the step of selecting a server load balancing type routing policy when said identified cache allocation approaches in value said identified trace footprint.

10. The method of claim 7, wherein said selecting step comprises the step of selecting a content localizing type routing policy when either said identified cache allocation is small, or when said trace footprint is large.

11. A machine readable storage having stored thereon a computer program for autonomic request routing policy selection, the computer program comprising a routine set of instructions which when executed by the machine cause the machine to perform the steps of:

- identifying a contemporary trace footprint experienced by a coupled server cluster;

- identifying a cache allocation for said coupled server cluster;

- retrieving at least two sets of hit rate metrics, each set of metrics corresponding to a particular routing policy;

- comparing said hit rate metrics based upon said identified trace footprint and said identified cache allocation to determine a preferred routing policy; and,

- selecting said preferred routing policy for use in routing content requests to said server cluster.

12. The machine readable storage of claim 11, further comprising the steps of:

computing with said hit rate metrics, an optimal server cluster configuration for said preferred routing policy; and,

provisioning an optimal number of servers in said server cluster based upon said computed optimal server cluster configuration.

13. The machine readable storage of claim 11, wherein said selecting step comprises the step of selecting a server load balancing type routing policy when said identified cache allocation approaches in value said identified trace footprint.

14. The machine readable storage of claim 11,, wherein said selecting step comprises the step of selecting a content localizing type routing policy when either said identified cache allocation is small, or when said trace footprint is large.